

Computing Laser Beam Paths in Optical Cavities:

An Approach based on Geometric Newton Method

Davide Cuccato · Alessandro Saccon ·

Antonello Ortolan · Alessandro Beghi

Received:date / Accepted: date

Davide Cuccato, Corresponding author

Department of Information Engineering, University of Padova

Via Gradenigo, 6 35100 Padova, Italy

cuccatod@dei.unipd.it

Alessandro Saccon

Department of Mechanical Engineering, Eindhoven University of Technology

De Rondon 70, 5612 AP, Eindhoven, the Netherlands

a.saccon@tue.nl

Antonello Ortolan

INFN - National Laboratories of Legnaro

Viale dell'Università, 2 35020 Legnaro (PD), Italy

antonello.ortolan@lnl.infn.it

Alessandro Beghi

Department of Information Engineering, University of Padova

Via Gradenigo, 6 35100 Padova, Italy

beghi@dei.unipd.it

Abstract In the last decade, increasing attention has been drawn to high precision optical experiments, which push resolution and accuracy of the measured quantities beyond their current limits. This challenge requires to place optical elements (e.g. mirrors, lenses, etc.) and to steer light beams with sub-nanometer precision. Existing methods for beam direction computing in resonators, e.g. iterative ray tracing or generalized ray transfer matrices, are either computationally expensive or rely on overparametrized models of optical elements. By exploiting Fermat’s principle, we develop a novel method to compute the steady-state beam configurations in resonant optical cavities formed by spherical mirrors, as a function of mirror positions and curvature radii. The proposed procedure is based on the geometric Newton method on matrix manifold, a tool with second order convergence rate that relies on a second order model of the cavity optical length. As we avoid coordinates to parametrize the beam position on mirror surfaces, the computation of the second order model does not involve the second derivatives of the parametrization. With the help of numerical tests, we show that the convergence properties of our procedure hold for non-planar polygonal cavities, and we assess the effectiveness of the geometric Newton method in determining their configurations with high degree of accuracy and negligible computational effort.

Keywords Geometric Newton method · Oblique manifold · Ring laser · Optical cavity

Mathematics Subject Classification (2000) 58E50, 49Q99

1 Introduction

In advanced optics applications, it is required to optimize the geometry of optical elements, in particular, when very high performance is sought for. One of such applications regards the design of resonant optical cavities that are essential elements of a wide range of devices and experiments, e.g. in laser physics, angular metrology, atomic clocks stabilization, etc. In this paper we focus on *ring laser gyroscopes*, which are devices used for measuring angular rotations with very high accuracy [1,2,3]. The core element of a ring laser is a three-dimensional resonant optical cavity, formed by $N > 2$ mirrors that are placed at the vertices of a polygon. In the cavity, the laser beam travels a closed optical path, defining a polygon of perimeter p and area \mathbf{a} [2]. In rotation sensing, \mathbf{a} and p are the most relevant geometric quantities, since their value and stability define the device performance in terms of sensitivity and accuracy.

Generally speaking, increasing the cavity dimensions (i.e. p and \mathbf{a}) results in measuring devices with higher sensitivity. In fact, the intrinsic noise limiting the sensitivity of a ring laser is the shot noise, and its magnitude turns out to be inversely proportional to p . However, the increase of dimensions negatively affects the ring laser long term stability, since changes of the environmental conditions (e.g. temperature and pressure drifts) during the measurement process induce geometry deformations which result in beam-jittering noise with magnitude almost proportional to p [1,4].

Even if a trade off can eventually be made between the intrinsic and beam-jittering noises, to increase sensitivity and stability of an optical cavity, the latter noise must be reduced as much as possible. To this aim, different approaches can be taken, leading to monolithic or heterolithic designs of the optical cavity. In the monolithic approach, one exploits an ultra low expansion material (e.g. Zerodur or Invar) to form a “rigid frame” supporting the mirrors, thus achieving passive stabilization of the cavity geometry by regulating pressure and temperature of the environment. For instance, the four-meter-wide square cavity of “G” [4] (presently the most sensitive and stable ring laser for geodetic and seismic applications) has a monolithic design. In the heterolithic design, mirrors are fixed to a concrete or granite frame and equipped with handlers to react against changes in their relative positions, thus stiffening the geometry of the apparatus [5,6]. Geometry can also be optimized to reduce the cavity sensitivity to the beam-jittering noise, e.g. by adjusting the beams path to regular polygonal shapes [7,8]. The heterolithic design overcomes the limitations due to the maximum size of a machinable monolithic element, and it is therefore chosen for very high performance applications, such as fundamental physics [9], geodesy and geophysics [10]. Clearly, implementation of the active geometry control of a heterolithic optical cavity requires the identification of suitable signals, provided e.g. by some metrological precision system, proportional to mirror displacements, and the derivations of accurate models.

In this paper we derive a geometrical model of an heterolithic ring laser to efficiently calculate the beams configuration as a function of the mirror

positions and orientations. This problem has already been addressed in the literature. In particular, generalized ray transfer matrices analysis, based on the optical axis perturbation, has been used in [11], whereas iterative ray tracing methods are used in [12]. These approaches are based on overparametrized models of the ring laser mirrors or involve a large number of iterations. To overcome these limitations, we exploit the Fermat's principle and the geometric Newton algorithm on matrix manifold. This tool has second order convergence rate and relies on a second order model of the objective function, that in the problem at hand is the cavity length p . In particular, the light path in a square cavity made of spherical mirrors is calculated, starting from the positions of their centers of curvature and the value of their curvature radii. As we avoid the use of coordinates to parametrize the beam position on mirror surfaces, the computation of the second order model does not involve the second derivatives of the parametrization. We show that the convergence properties of our procedure hold for the optical cavities of interest. Finally, we assess the effectiveness of the geometric Newton method in determining their configurations with high degree of accuracy and negligible computational effort.

The paper is organized as follows. In Sect. 2 notations and definitions are given. Sect. 3 is devoted to the problem statement and formulation. In Sect. 4 we review the geometric Newton Algorithm, which is then specialized in Sect. 5 to the Oblique Manifold. In Sect. 6 the application of the proposed algorithm to the square ring laser cavity is presented, and in Sect. 7 some numerical results are presented. Conclusions are drawn in Sect. 8.

2 Notation and Mathematical Preliminaries

In this paper we make use of the theory of finite dimensional smooth manifolds and covariant differentiation as presented in [13,14]. The symbols in Tab. 2 will be used throughout the paper.

Notation

\mathcal{E}	Euclidean space \mathcal{E} .
$\mathcal{M}, \mathcal{N} \subseteq \mathcal{E}$	Embedded Submanifolds $\mathcal{M}, \mathcal{N} \subseteq \mathcal{E}$.
$x \in \mathcal{M}$	Element x of the manifold \mathcal{M} .
$f : \mathcal{M} \rightarrow \mathbb{R}, \bar{f} : \mathcal{E} \rightarrow \mathbb{R}$	Real valued functions on \mathcal{M} and \mathcal{E} such that $\bar{f}(x) = f(x)$ for $x \in \mathcal{M}$.
$\mathfrak{F}(\mathcal{M})$	The set of smooth real valued functions on \mathcal{M} .
$\mathfrak{F}_x(\mathcal{M})$	The set of smooth real valued function defined near $x \in \mathcal{M}$.
$T_x \mathcal{M}$	The tangent space to \mathcal{M} at $x \in \mathcal{M}$.
$\xi_x \in T_x \mathcal{M}$	The tangent vector ξ_x to \mathcal{M} at x .
$\mathfrak{X}_x(\mathcal{M})$	The set of smooth vector fields on \mathcal{M} near x .
$\xi \in \mathfrak{X}_x(\mathcal{M})$	Smooth vector field $\xi : x \mapsto \xi_x$ on \mathcal{M} at x .
$DF : T_x \mathcal{M} \rightarrow T_{F(x)} \mathcal{N}$	The tangent map of $F : \mathcal{M} \rightarrow \mathcal{N}$ at x .
$\partial f(x) \in T_x \mathcal{E}$	Euclidean Gradient of $f : \mathcal{E} \rightarrow \mathbb{R}$.
$\partial^2 f(x) : T_x \mathcal{E} \rightarrow T_x \mathcal{E}$	Euclidean Hessian of f .
$\langle \cdot, \cdot \rangle_{\mathcal{M}} : T_x \mathcal{M} \times T_x \mathcal{M} \rightarrow \mathbb{R}$	Riemannian metric on \mathcal{M} .
$\nabla : T_x \mathcal{M} \times \mathfrak{X}_x(\mathcal{M}) \rightarrow \mathfrak{X}_x(\mathcal{M})$	Riemannian connection on $\mathcal{M} \subset \mathcal{E}$.
$\text{grad } f(x) \in T_x \mathcal{M}$	Riemannian Gradient of $f : \mathcal{M} \rightarrow \mathbb{R}$.
$\text{Hess } f(x) : T_x \mathcal{M} \rightarrow T_x \mathcal{M}$	Riemannian Hessian of f .
\mathbb{S}^2	Unit Sphere $\{\mathbf{x} \in \mathbb{R}^3, \mathbf{x}^T \mathbf{x} = 1\}$
\wedge	Canonical cross product between vectors of \mathbb{R}^3 .
\otimes	Kronecker product of two matrices,
$\text{diag} : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$	$[\text{diag}(A)]_{ij} = \begin{cases} A_{ij} & \text{if } i = j, \\ 0 & \text{otherwise} \end{cases}$

The Riemannian connection ∇ is determined by the condition

$$D\langle \xi_x, \chi_x \rangle_{\mathcal{M}}[\eta_x] = \langle (\nabla_{\eta_x} \xi)_x, \chi_x \rangle_{\mathcal{M}} + \langle \xi_x, (\nabla_{\eta_x} \chi)_x \rangle_{\mathcal{M}}, \quad (2.1)$$

where $x \in \mathcal{M}$, $\eta_x \in T_x \mathcal{M}$, and $\xi, \chi \in \mathfrak{X}_x(\mathcal{M})$, i.e. it is the unique connection that is compatible with the Riemannian metric $\langle \cdot, \cdot \rangle_{\mathcal{M}}$ of \mathcal{M} . Moreover, if $\mathcal{M} = \mathcal{E}$, we have that the covariant derivative associated to the Riemannian connection of the vector field ξ with respect to η_x is simply the directional derivative, i.e. $\nabla_{\eta_x} \xi = D\xi[\eta_x]$.

3 Problem Statement and Formulation

We consider an optical cavity formed by N spherical mirrors. By indicating with $\mathbf{z}_k \in \mathbb{R}^3$ the coordinates of the position of the light spot on the k -th mirror with respect to the ground frame, the cavity optical length can be computed as $p = \sum_{k=1}^N \|\mathbf{z}_k - \mathbf{z}_{k+1}\|$ and the associated vector area (with magnitude equal to the area of the cavity and direction perpendicular to the cavity plane) as $\mathbf{a} = \frac{1}{2} \sum_{k=1}^N \mathbf{z}_k \wedge \mathbf{z}_{k+1}$, where we pose $\mathbf{z}_{N+1} := \mathbf{z}_1$. By applying the formalism of geometric optics, we can model the k -th spherical mirror M_k as a sphere of center $\mathbf{c}_k \in \mathbb{R}^3$ and curvature radius $r_k \in \mathbb{R}^+$. The position of the laser spot on the k -th mirror can be expressed as $\mathbf{z}_k = \mathbf{c}_k + r_k \mathbf{x}_k$, where $\mathbf{x}_k \in \mathbb{S}^2$, see Figure 3.1.

A *configuration* for the laser beams in the optical cavity is defined as the ordered set of points $\mathcal{X} = (\mathbf{x}_1, \dots, \mathbf{x}_N) \in \underbrace{\mathbb{S}^2 \times \dots \times \mathbb{S}^2}_{N \text{ times}}$, which describe the

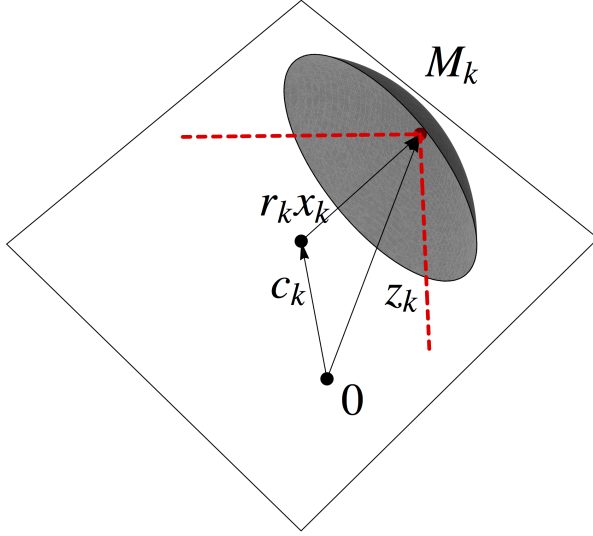


Fig. 3.1 Parametrization of the position of the laser spot on the k -th mirror M_k . With respect to the inertial frame origin O , the position of laser spot is $\mathbf{z}_k = r_k \mathbf{x}_k + \mathbf{c}_k$. The mirror M_k has radius r_k . Relative to the mirror center \mathbf{c}_k , the position of the laser spot can be parametrized as $r_k \mathbf{x}_k$, $\mathbf{x}_k \in \mathbb{S}^2$.

positions of laser light spots on each of the N mirrors. In addition, we define the matrix of centers $\mathcal{C} = (\mathbf{c}_1, \dots, \mathbf{c}_N)$ and the matrix of curvature radii $\mathcal{R} = \text{diag}(r_1, \dots, r_N)$. We refer to \mathcal{R} and \mathcal{C} as the *parameters* of the optical cavity. Given the cavity parameters \mathcal{R} and \mathcal{C} and a configuration \mathcal{X} , the corresponding laser beams path as well as the associated (scalar and vector) fields p and \mathbf{a} (see Section 1) can be computed.

Note that not every configuration \mathcal{X} is admissible as physical solution. In fact, the physical light paths have to be stationary points of p with respect to \mathcal{X} [15, Chap. 2]. This fact results from the application of the Hamilton's principle to geometric optics, when we identify the Hamiltonian coordinates

with the Euclidean coordinates of the laser spots and the action functional with the light optical path length. This result is known in optics as the Fermat's principle, stating that the admissible paths are the ones that make their length stationary with respect to infinitesimal variations of the paths.

We turn now to the problem of computing the physical beam configuration given the cavity parameters \mathcal{R} and \mathcal{C} . To this end, we resort to an intrinsic geometric approach that avoids a local parametrization of the configuration manifold, only constraining the laser spots on each mirror to lie on the surface of a sphere. The advantages of this approach are to handle simpler expressions and to avoid the use of multidimensional spherical coordinates and derivatives of the parametrization.

The optical path length is a function of the cavity configuration \mathcal{X} and of the cavity parameters \mathcal{C} and \mathcal{R} . The configuration \mathcal{X} lives in the cartesian product of N unit spheres in \mathbb{R}^3 , $\underbrace{\mathbb{S}^2 \times \dots \times \mathbb{S}^2}_{N \text{ times}}$. It is straightforward to show that this configuration manifold is an embedded submanifold of $\mathbb{R}^{3 \times N}$. This manifold is usually referred to as the Oblique Manifold of dimension $2 \times N$ [13]. The Oblique Manifold of dimension $n \times m$ is defined as

$$\mathcal{OB}(n, m) = \left\{ \mathcal{X} \in \mathbb{R}^{(n+1) \times m}, d(\mathcal{X}^T \mathcal{X}) = I_{m \times m} \right\} \cong \underbrace{\mathbb{S}^n \times \dots \times \mathbb{S}^n}_m, \quad (3.1)$$

where the function d is defined in Tab. 2, and the tangent space at \mathcal{X} of $\mathcal{OB}(n, m)$ is

$$T_{\mathcal{X}}\mathcal{OB}(n, m) = \left\{ Y \in \mathbb{R}^{(n+1) \times m} : d(\mathcal{X}^T Y) = 0_{m \times m} \right\}. \quad (3.2)$$

It is worth mentioning that $\mathcal{OB}(n, m)$ inherits the (Riemannian) metric of $\mathbb{R}^{n \times m}$, i.e., we can define its Riemannian metric simply as $\langle \mathcal{X}, \mathcal{Y} \rangle_{\mathcal{OB}(n, m)} = \text{tr}(\mathcal{X}^T \mathcal{Y})$. Moreover, each column of \mathcal{X} in (3.2) is orthogonal to the corresponding column of \mathcal{Y} , with respect to the (Riemannian) metric of \mathbb{R}^{n+1} , $\langle \mathbf{x}, \mathbf{y} \rangle_{\mathbb{R}^{n+1}} = \mathbf{x}^T \mathbf{y}$. In fact the columns of \mathcal{X} have the straightforward interpretation of points on \mathbb{S}^2 , so that each column of \mathcal{Y} represents a possible tangent vector to the corresponding column of \mathcal{X} , seen as a point of \mathbb{S}^2 .

A *physical configuration* $\hat{\mathcal{X}}$ is a cavity configuration that also satisfies Fermat's principle, such that it is a stationary point of the optical cavity length $p = p(\mathcal{X}; \mathcal{C}, \mathcal{R})$. Formally, the set of physical configurations is given by

$$\{\mathcal{X} \in \mathcal{OB}(2, N) : T_{\mathcal{X}}\mathcal{OB}(2, N) \ni \text{grad } p(\mathcal{X}; \mathcal{C}, \mathcal{R}) = 0\}. \quad (3.3)$$

By the Weierstrass theorem the set (3.3) can contain 2 or more elements being p a continuous function defined over a compact set $\mathcal{OB}(2, N)$. It is not possible, in general, to find closed form expressions for the elements of the set of physical configurations, therefore we resort to a numerical algorithm. We consider optical cavities with parameters \mathcal{C} and \mathcal{R} which slightly differ from the nominal values $\mathcal{C}^*, \mathcal{R}^*$. We assume also that the physical configuration \mathcal{X}^* of the optical cavity with parameters $\mathcal{C}^*, \mathcal{R}^*$ is known. Then, the physical configuration \mathcal{X}^* can be used as the initial condition for our algorithm. To efficiently compute a solution with desired accuracy, we propose to use a geometric Newton's method, which requires the computation of first and second order derivatives of the functions $p : \mathcal{OB}(2, N) \mapsto \mathbb{R}$. Such computation can be carried out in an elegant and efficient way by using advanced tools from

differential geometry, involving the curvature and affine connection associated with the Riemannian manifold $\mathcal{OB}(2, N)$.

4 Stationary points of a vector field on a Riemannian manifold

As introduced in the previous section, our aim is to create an algorithm to find physical configurations \mathcal{X} of an optical cavity whose geometry is characterized by the mirror center positions \mathcal{C} and curvature radii \mathcal{R} . We model the configuration space as the Riemannian manifold $\mathcal{OB}(2, N)$ and we call $p : \mathcal{OB}(2, N) \rightarrow \mathbb{R}$ the length of the light path associated to a given configuration \mathcal{X} . In this setting, Fermat's principle is equivalent to finding the *stationary* points of the vector field $\text{grad } p$ (a physical configuration \mathcal{X} is, in general, just a extremal point of the light-path length p , and not a minimizer). This section provides a review on an efficient Newton method that can be used for finding stationary points of a vector field on a Riemannian manifold, *in the specific situation where this Riemannian manifold is embedded into an Euclidean space*.

This section details the algorithm in its general form, considering an arbitrary smooth Riemannian manifold \mathcal{M} endowed with a Riemannian metric $\langle \cdot, \cdot \rangle_{\mathcal{M}}$ and a real valued function $f \in \mathfrak{F}(\mathcal{M})$ defined on \mathcal{M} embedded in the Euclidean space \mathcal{E} . The specific case $f = p$, $\mathcal{M} = \mathcal{OB}(2, N)$, and $\mathcal{E} = \mathbb{R}^{3 \times N}$ will be treated in the next section.

The extensions of Newton method to find stationary points of a function defined on an embedded submanifold of the Euclidean space \mathcal{E} (this trivial

manifold corresponding to \mathbb{R}^n with the standard norm) has received a substantial amount of attention in the recent years as the exploitation of geometric properties leads to effective, elegant, and efficient algorithms [16, 17, 18, 19, 20]. We follow here the approach as outlined in [13, 21, 22]. The typical way of approaching the problem of finding the stationary points of a vector field is to reformulate the root finding problem $\text{grad } p(x) = 0$, $x \in \mathcal{M}$ into the minimization problem $\min h(x) := |\text{grad } p(x)|^2$, $x \in \mathcal{M}$. Naively, one could then employ Newton method to minimize the newly introduced function h over \mathcal{M} by computing a second order approximation of h based on its gradient and Hessian. This naive approach has the disadvantage of requiring the third order derivatives of the length function p for the computation of the Hessian of h . However, an efficient alternative method exists that makes use of both functions h and p , just requiring the computation of the gradient and Hessian of p . Such algorithm, based on Newton method, exhibits second order convergence rate in a neighborhood of stationary point as long as the Hessian of p is not degenerate. As shown in [13], at each iteration a line search algorithm (e.g., Armijo's [23]) is run for the function h along a descent direction that is computed via the gradient and Hessian of p .

4.1 The algorithm

A key ingredient in geometric optimization is the use of a *retraction* [13, Ch.4].

A retraction $R_x : T_x \mathcal{M} \mapsto \mathcal{M}$ allows one to use the tangent space $T_x \mathcal{M}$ at a given point $x \in \mathcal{M}$ as a local parametrization of the neighborhood of x .

Retractions are used in the update step of the geometric algorithm, when a descent direction is turned into the next iterate point. The ideal retraction map is the exponential map associated to the Riemannian connection of \mathcal{M} . As computing the exponential map is sometimes prohibitive or time consuming, suitable approximations of the exponential map (that agree up to the second derivative at the origin with the exponential map) can be used, still retaining the second order convergence rate of the algorithm. When the manifold is embedded in an Euclidean space, there is a standard method to obtain a retraction map based on tangent space projection.

The geometric Newton algorithm for finding a stationary point of a vector field on a Riemannian manifold can be described at high level as follows.

Algorithm 1 *Geometric Newton method for finding stationary points*

Input: $x_0 \in \mathcal{M}$, real valued function f on \mathcal{M}

Output: Sequence of iterates x_1, x_2, \dots

1. [Search direction] Solve (4.1) to get the descent direction $\eta_{x_k} \in T_{x_k} \mathcal{M}$.
 2. [Line Search] Find t_k that approximately solves $\arg \min_{\lambda} h(R_x(\lambda \eta_{x_k}))$
 3. [Update] Set $x_{k+1} = R_x(t_k \eta_{x_k})$
-

The search direction computation requires the solution of the classical geometric Newton equation (cf. [13, Ch.6])

$$\text{Hess } f(x_k)[\eta_{x_k}] = -\text{grad } f(x_k), \quad \eta_{x_k} \in T_{x_k} \mathcal{M}, \quad (4.1)$$

where the computation of the Riemannian Hessian requires the Riemannian connection ∇ associated with M . The search direction computation corre-

sponds to the minimization of the quadratic model \tilde{f} of f centered at the current iterate $x_k \in \mathcal{M}$ defined as

$$\tilde{f}(x_k, \eta_{x_k}) = f(x_k) + \langle \text{grad } f(x_k), \eta_{x_k} \rangle + \frac{1}{2} \langle \text{Hess } f(x_k)[\eta_{x_k}], \eta_{x_k} \rangle. \quad (4.2)$$

After the computation of the descent direction $\eta_{x_k} \in T_{x_k} \mathcal{M}$, a line search algorithm is employed, now on the function h . An effective line search algorithm is the Armijo's [24]. At each iterate, the step size t is set to $\alpha \beta^l$ (backtracking approach), with l the smallest integer such that

$$h(R(t \eta_x)) \leq h(x) + \sigma t Dh(x)[\eta_x], \quad (4.3)$$

with $x = x_k$ denoting the current iterate, $\eta_x = \eta_{x_k}$ the current descent direction, and $\alpha > 0$, and $\beta \in (0, 1)$, and $\sigma \in (0, 1)$ design parameters. The condition (4.3) assure the convergence of the line search if the function $h(R(t \eta_x))$ to be minimized is sufficiently smooth (for a proof, see, e.g., [13, Chapter 4]).

Once the search direction and the step size have been found, the next iterate is computed and the procedure is repeated as long as a stop criterium is not met (norm of $\text{grad } f$ sufficiently small). The proof of convergence of Algorithm 1 to a stationary point of $\text{grad } f$ can be obtained by a simple adaption of the convergence result presented in [13, Chapter 4] and is described in the Appendix.

4.2 The exploitation of the embedding of \mathcal{M} into \mathcal{E}

In the following, we detail how to solve (4.1) exploiting the fact that \mathcal{M} is embedded in the Euclidean Space \mathcal{E} . In particular, such a hypothesis allows

treating both points of \mathcal{M} and vectors of $T\mathcal{M}$ as elements of $\mathcal{E} \simeq T\mathcal{E}$ and use the trivial Riemannian connection of \mathcal{E} to perform covariant differentiation on \mathcal{M} to obtain the Riemannian Hessian of f .

Given a point $x \in \mathcal{M}$, the Newton equation requires the solution of the linear system (4.1) to find the descent direction η_x . We recall that, on a manifold, the Riemannian gradient and Hessian at $x \in \mathcal{M}$ are defined such that

$$\langle \text{grad } f(x), \xi_x \rangle_{\mathcal{M}} = Df(x)[\xi_x] \quad (4.4)$$

$$\text{Hess } f(x)[\eta_x] = (\nabla_{\eta_x} \text{grad } f)_x, \quad (4.5)$$

where Df denote the differential of f and ∇ covariant differentiation associated to the Riemannian metric of \mathcal{M} . When \mathcal{M} is an embedding of \mathcal{E} , since we can represent vectors in $T_x\mathcal{E}$ as points in \mathcal{E} ($T_x\mathcal{E} \simeq \mathcal{E}$) and $T_x\mathcal{M} \subseteq T_x\mathcal{E}$, we can represent points in \mathcal{M} and \mathcal{E} and tangent vectors in $T_x\mathcal{M}$ and $T_x\mathcal{E}$ using the same representation [13, Chapter 3].

Computation of the Riemannian gradient and Hessian of a function f defined on \mathcal{M} is achieved by computing the standard gradient and Hessian of any smooth extension \bar{f} of f (a smooth function whose restriction on M corresponds to f) and employing the tangent space projection detailed below. The gradient and Hessian of \bar{f} in \mathcal{E} will be denoted $\partial\bar{f}$ and $\partial^2\bar{f}$ respectively, and called, from now on, Euclidean gradient and Euclidean Hessian of \bar{f} . The interested reader is referred to [13, Chapter 5] for further details about the proof of the following results.

Tangent Space Projection. For any $x \in \mathcal{M}$, the orthogonal projection operator

$$P_x : T_x \mathcal{E} \rightarrow T_x \mathcal{M} \quad (4.6)$$

$$\xi_x \mapsto P_x(\xi_x)$$

maps an arbitrary tangent vector ξ_x of $T_x \mathcal{E}$ into a vector of $T_x \mathcal{M}$ corresponding to the orthogonal projection of $T_x \mathcal{E}$ into its subspace $T_x \mathcal{M}$. In addition, every tangent vector ξ_x in $T_x \mathcal{E}$ can be decomposed as the direct sum $\xi_x = \eta_x + \nu_x$, where $\eta_x \in T_x \mathcal{M}$ and $\nu_x \in (T_x \mathcal{M})^\perp$, the normal space of \mathcal{M} at x , i.e. the orthogonal complement of $T_x \mathcal{M}$ in $T_x \mathcal{E}$. It follows that if $\eta_x = P_x(\xi_x)$ then $\nu_x = \xi_x - P_x(\xi_x)$ [13].

Riemannian gradient. The Riemannian gradient $\text{grad } f$ is computed via the Euclidean gradient $\partial \bar{f}$ and the tangent space projection (4.6) as

$$\text{grad } f(x) = P_x(\partial \bar{f}(x)), \quad (4.7)$$

for every $x \in \mathcal{M}$.

Riemannian Hessian. Denote with

$$D \text{grad } \bar{f}(x) : T_x \mathcal{E} \mapsto T_x \mathcal{E} \quad (4.8)$$

the tangent map of $\text{grad } f$ defined in (4.7) seen as the function $\bar{P}_x(\partial \bar{f}(\cdot)) : \mathcal{E} \rightarrow \mathcal{E}$, using the identification $T_x \mathcal{E} \simeq \mathcal{E}$ and where $\bar{P}_x : T_x \mathcal{E} \rightarrow T_x \mathcal{E}$ is a smooth extension of P_x defined for $x \in \mathcal{E}$. Then, the Riemannian Hessian of f at x can be computed as

$$\text{Hess } f(x)[\eta] = P_x(D \text{grad } \bar{f}(x)[\eta]) \quad (4.9)$$

for all $x \in \mathcal{M}$. We recall that, in general, $D \operatorname{grad} \bar{f}(x) \neq D \partial \bar{f}(x) = \partial^2 \bar{f}(x)$.

It is not straightforward to give a general expression for $D \operatorname{grad} \bar{f}$ in (4.7). Usually, this relation is expressed in terms of the Weingarten map, also known as shape operator[21]. In this paper, we limit ourself to provide the specific expression for $D \operatorname{grad} \bar{f}(x)$ for $\mathcal{M} = \mathcal{OB}(2, N)$.

4.3 The special case $\mathcal{M} = \mathcal{OB}(2, N)$ and $\mathcal{E} = \mathbb{R}^{3 \times N}$

We restrict here our analysis to scalar fields defined on $\mathcal{OB}(2, N)$. As the oblique manifold $\mathcal{OB}(2, N)$ can be embedded in the Euclidean space \mathcal{E} of dimension $3 \times N$, a point $x \in \mathcal{OB}(2, N)$ and a tangent vector $\xi_x \in T_x \mathcal{OB}(2, N)$ can both be represented by vectors in $\mathbb{R}^{3 \times N}$. We detail in the following how to compute the gradient, Hessian, and retraction required in Algorithm 1 for the specific case $\mathcal{M} = \mathcal{OB}(2, N)$, relevant to our application.

As we saw in the previous section (in particular, (4.7) and (4.9)), the computation of the Riemannian gradient and Hessian requires the computation of the orthogonal projection $P_x(\xi_x)$ with $\xi_x \in T_x \mathcal{M}$ and the directional derivative $D \operatorname{grad} \bar{f}(x)$. Indicating with $X \in \mathbb{R}^{3 \times N}$ the (redundant) parameterization of a point $x \in \mathcal{OB}(2, N)$ and with $\xi_X \in \mathbb{R}^{3 \times N}$ the parametrization of a vector $\xi_x \in T_x \mathcal{OB}(2, N)$, these two linear operators can be computed by the expressions provided in Table 1. Note in particular that $\partial \bar{f}(\mathbf{X})$ and $\partial^2 \bar{f}(\mathbf{X})[\eta_{\mathbf{X}}]$ are $3 \times N$ matrices. The formulas are obtained by straightforward (matrix) differentiation. Recall that the function $\operatorname{diag} : \mathbb{R}^{N \times N} \rightarrow \mathbb{R}^{N \times N}$, defined in the notation section, simply extracts the diagonal matrix from a given matrix.

Note furthermore that the term $X \operatorname{diag}(\eta_X^T \partial \bar{f}(X))$ appearing in the expression for $D \operatorname{grad} \bar{f}(X)[\eta_X]$ defined in Table 1 has no influence in the computation of the Hessian as its projection is zero: its computation can therefore be avoided to save computational time. Further details on how these formulas have been obtained can be found in [22]

Abstract terms	Coordinate representation in $\mathbb{R}^{3 \times N}$
x	$X \in \mathbb{R}^{3 \times N}$
ξ_x	$\xi_X \in \mathbb{R}^{3 \times N}$
$P_x(\xi_x)$	$P_X(\xi_X) := \xi_X - X \operatorname{diag}(X^T \xi_X) \in \mathbb{R}^{3 \times N}$
$D \operatorname{grad} \bar{f}(x)[\eta_x]$	$D \operatorname{grad} \bar{f}(X)[\eta_X] := P_X(\partial^2 \bar{f}(X)[\eta_X]) - \eta_X \operatorname{diag}(X^T \partial \bar{f}(X)) - X \operatorname{diag}(\eta_X^T \partial \bar{f}(X)) \in \mathbb{R}^{3 \times N}$

Table 1 Coordinate representation for $P_x(\xi_x)$ and $D \operatorname{grad} \bar{f}(x)[\eta_x]$ for the $\mathcal{OB}(2, N)$.

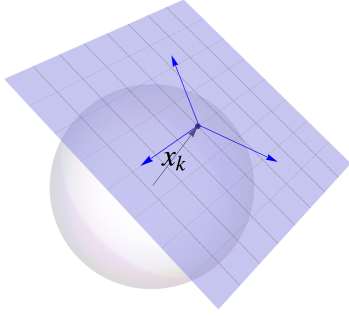
Solution to the geometric Newton equation.

The geometric Newton equation (4.1) is a linear equation with solution $\eta_X \in T_X \mathcal{OB}(2, N)$. Here we detail how to solve this abstract problem by choosing, for a given configuration $X \in \mathcal{OB}(2, N) \subset \mathbb{R}^{3 \times N}$, a suitable basis for the tangent space $T_X \mathcal{OB}(2, N) \subset T_X \mathbb{R}^{3 \times N}$, turning the linear equation into a classical set of linear equations for the coordinate representation of the “descent” direction η_X .

When considering $\mathcal{OB}(2, N)$ as embedded in $\mathbb{R}^{3 \times N}$, a point in $X \in \mathcal{OB}(2, N) \subset \mathbb{R}^{3 \times N}$ can be written simply as the matrix $X = [\mathbf{x}_1, \dots, \mathbf{x}_N]$, with $\mathbf{x}_i \in \mathbb{S}^2 \subset \mathbb{R}^3$. Recall that $\mathcal{OB}(2, N)$ and, consequently, $T_X \mathcal{OB}(2, N)$ for a given $X \in \mathcal{OB}(2, N)$ have dimension $2N$. Define $\pi(\mathbf{x}) := I_3 - \mathbf{x}\mathbf{x}^T$, $\mathbf{x} \in \mathbb{S}^2$, the

orthogonal projection of a vector in \mathbb{R}^3 into the 2-dimensional subspace of \mathbb{R}^3 that is orthogonal to the unit vector \mathbf{x} . All rows of the matrix $\pi(\mathbf{x})$ are orthogonal to the unit vector \mathbf{x} (for some \mathbf{x} , one of the row of $\pi(\mathbf{x})$ can be zero). To obtain a basis for $T_{\mathbf{x}}\mathbb{S}^2$, one can pick (one of) the row of $\pi(\mathbf{x})$ with maximum norm and identify it as a vector \mathbf{v} that one uses as the first element of the basis. One then defines $\mathbf{w} := \mathbf{v} \times \mathbf{x}$ as the second element of the basis. In general, \mathbf{v} and \mathbf{w} are not unit vectors. The selection of \mathbf{v} and \mathbf{w} as a function of \mathbf{x} cannot be continuous over the entire sphere (this is a consequence of the hairy ball theorem), but the lack of continuity is not an issue as one just need a basis of the tangent space to solve the Newton equation (4.1). A simple visualization of the just mentioned *project-and-select* approach is given in Figure 4.1.

a)



b)

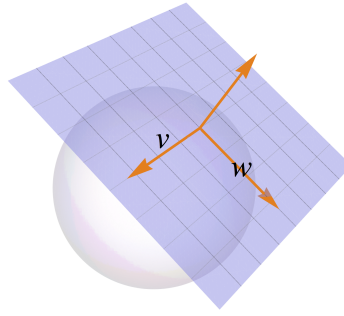


Fig. 4.1 a) the point \mathbf{x}_i , its tangent space $T_{\mathbf{x}_i}(\mathbb{S}^2)$, and the three row vectors of the matrix which represents $\pi(\mathbf{x}_i)$; b) the tangent space $T_{\mathbf{x}_i}(\mathbb{S}^2)$ and the vectors v and w .

As $\mathcal{OB}(2, N)$ is simply the cartesian product of N copies of the unit sphere \mathbb{S}^2 , we can use the same project-and-select approach to construct a basis for $T_{\mathbf{X}}(\mathcal{OB}(2, N)) \subset T_X \mathbb{R}^{3 \times N}$, $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_N]$, by combining N bases obtained with the project-and-select approach for the spaces $T_{\mathbf{x}_i} \mathbb{S}^2$. Formally, denoting \mathbf{v}_i and $\mathbf{w}_j \in \mathbb{R}^3$ the basis for $T_{\mathbf{x}_i} \mathbb{S}^2$, a basis for $T_{\mathbf{X}} \mathcal{OB}(2, N) \subset \mathbb{R}^{3 \times N}$ is the set of $2N$ matrices in $\mathbb{R}^{3 \times N}$ defined as $\bigcup_{i=1}^N \left\{ \mathbf{v}_i \otimes \mathbf{e}_i^T, \mathbf{w}_i \otimes \mathbf{e}_i^T \right\}$, where $\mathbf{e}_i \in \mathbb{R}^N$ denotes the standard i -th unit vector of \mathbb{R}^N and \otimes the Kronecker product. We immediately also get a (nonstandard) basis for $T_{\mathbf{X}} \mathbb{R}^3 = T_{\mathbf{X}} \mathcal{OB}(2, N) \oplus T_{\mathbf{X}}^\perp \mathcal{OB}(2, N)$ simply adding the orthogonal spaces spanned by \mathbf{x}_i as $\bigcup_{i=1}^N \left\{ \mathbf{x}_i \otimes \mathbf{e}_i^T, \mathbf{v}_i \otimes \mathbf{e}_i^T, \mathbf{w}_i \otimes \mathbf{e}_i^T \right\}$.

Proposition 4.1 *Let $X \in \mathcal{OB}(2, N)$ and $f : \mathcal{OB}(2, N) \rightarrow \mathbb{R}$ with extension $\bar{f} : \mathbb{R}^{3 \times N} \rightarrow \mathbb{R}$ (i.e., $\bar{f}|_{\mathcal{OB}(2, N)} = f$). Let $\mathbf{b}_k(X)$, $k = \{1, 2, \dots, 2N\}$ denote the basis of $T_X \mathcal{OB}(2, N)$ using the project-and-select approach, then the Newton equation (4.1) for f with respect to this basis reads*

$$H(X)\mathbf{y} = \mathbf{g}(X) \quad (4.10)$$

with $\mathbf{y} \in \mathbb{R}^{2N}$ is the coordinate representation of η_X , $\eta_X = \mathbf{y}_j \mathbf{b}_j(X)$, and the matrix $H(X) \in \mathbb{R}^{2N \times 2N}$ and $\mathbf{g}(X) \in \mathbb{R}^{2N}$ satisfy

$$[g(X)]_k \mathbf{b}_k(X) = \text{grad } f(X) = P_X(\partial \bar{f}(X)) \quad (4.11)$$

$$[H(X)]_{ij} \mathbf{b}_i(X) = \text{Hess} f(X)[\mathbf{b}_j(X)] = P_X(D \text{grad } \bar{f}(X)[\mathbf{b}_j(X)]) \quad (4.12)$$

where P_X and $D \text{grad } \bar{f}(X)$ in $\mathcal{OB}(2, N)$ as given in in Table 1.

The proof of Proposition 4.1 is a straightforward exercise based on employing the definition of the basis $\mathbf{b}_k(X)$. Note that the ordering of the basis is com-

pletely arbitrary: In our implementation, we used $\mathbf{v}_1 \otimes \mathbf{e}_1^T, \mathbf{v}_2 \otimes \mathbf{e}_2^T, \dots, \mathbf{v}_N \otimes \mathbf{e}_N^T$, $\mathbf{w}_1 \otimes \mathbf{e}_1^T, \mathbf{w}_2 \otimes \mathbf{e}_2^T, \dots, \mathbf{w}_N \otimes \mathbf{e}_N^T$, but any other choice would be legitimate.

A retraction for $\mathcal{OB}(2, N)$.

Despite the fact that for each Riemannian manifold, the exponential map is a natural retraction, different retractions are usually employed to minimize computational costs without compromising the convergence rate of the zero finding (or minimizing) algorithms. For $\mathcal{OB}(2, N)$, in particular, the map

$$R(X, \xi_X) = (X + \xi_X) \operatorname{diag} \left((X + \xi_X)^T (X + \xi_X) \right)^{-1/2}, \quad (4.13)$$

can be used. Note that the square root and inversion operation has lower priority than the diag operator (defined in notation section). Geometrically, this retraction corresponds to the normalization to unit norm of each column of the matrix $X + \xi_X$ to bring it back to $\mathcal{OB}(2, N)$ [13, 21, 22].

5 Physical configurations of a ring laser cavity: The algorithm

In the previous section, we have detailed how the gradient and Hessian of a function defined on $\mathcal{OB}(2, N)$ can be represented and used in solving the Newton equation (4.1) for a generic function defined on $\mathcal{OB}(2, N)$. This section shows how these results can be applied to obtain an algorithm to compute the physical configuration of a ring laser cavity with four spherical mirrors. From now on, therefore, $\mathcal{E} = \mathbb{R}^{3 \times 4}$ and $\mathcal{M} = \mathcal{OB}(2, 4)$.

In the following, we detail how to compute the function $p : \mathcal{OB}(2, 4) \rightarrow \mathbb{R}$ representing the laser path length. Recall that our goal is find the extremal points of this function as they represent the physical configurations consistent with Fermat's principle.

The configuration of the optical cavity is

$$X = (\mathbf{x}_1, \dots, \mathbf{x}_4) \in \mathcal{OB}(2, 4), \quad (5.1)$$

while its parameters, namely the mirror centers and radii, are given, respectively, by

$$C = (\mathbf{c}_1, \dots, \mathbf{c}_4) \in \mathbb{R}^{3 \times 4}, \quad (5.2)$$

$$Q = \text{diagonal}(r_1, r_2, r_3, r_4) \in \mathbb{R}^{4 \times 4}. \quad (5.3)$$

In matrix form, the coordinates of the light spots on the mirror surface are given by

$$Z(X) := XQ + C, \quad (5.4)$$

where $Z = (\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3, \mathbf{z}_4) \in \mathbb{R}^{3 \times 4}$. Figure 5.1 provides a graphical representation of a typical disposition for the mirror centers, light spots, and laser beam path.

The vectors joining consecutive light spots on the mirror surfaces, describing the sides of the polygonal cavity, are given by $\mathbf{y}_k = (\mathbf{z}_{k+1} - \mathbf{z}_k)$, $k = \{1, 2, 3, 4\}$, with the convention that $\mathbf{z}_5 := \mathbf{z}_1$. The length of the optical

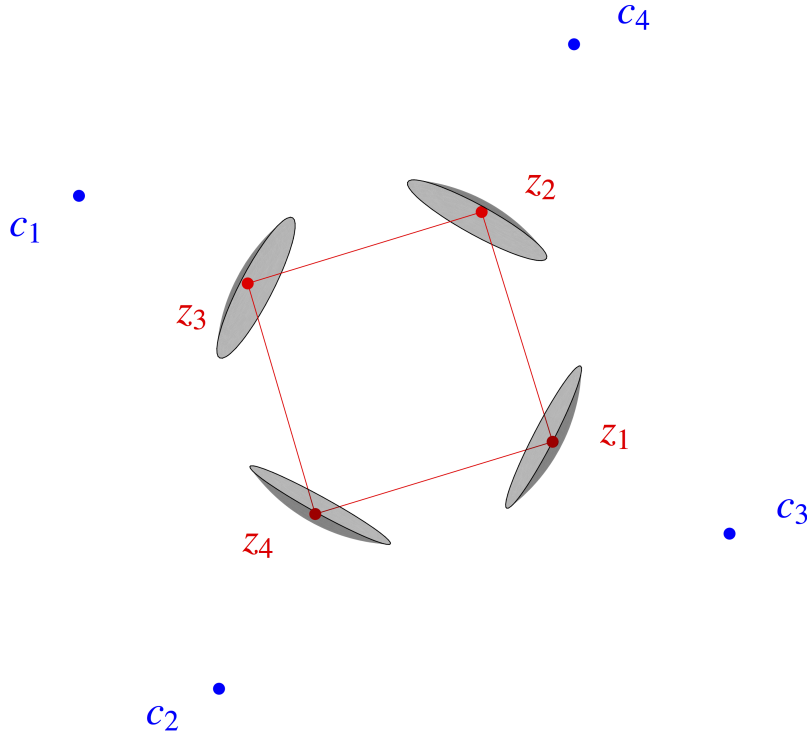


Fig. 5.1 An example of displacements of mirrors forming a square optical cavity. Elements relative to each mirror are colored in grey. Red dots mark the mirror centers, the red lines represent the optical path length as a link between consecutive laser spots z_k .

path is therefore $p(X; C, Q) = \sum_{i=1}^4 \|y_i\|$. Defining $Y(X) := Z(X)M$ with

$$M = \begin{bmatrix} 1 & 0 & 0 & -1 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix},$$

the optical path length can be written, in matrix form, as

$$p(X; C, Q) = \text{tr} \left(\text{diag} \left(Y^T(X) Y(X) \right)^{1/2} \right). \quad (5.5)$$

This is the formula that can be used to derive the Netwon method on $\mathcal{OB}(2, 4)$.

Note that in (5.5), the square root operator $(\cdot)^{1/2}$ acts component wise on the entries of the diagonal matrix and tr denotes the trace operator.

Equation (5.5) also defines the extension \bar{p} of p to $\mathcal{E} = \mathbb{R}^{3 \times 4}$. The Euclidian gradient and Hessian of \bar{p} , appearing in the formulas presented in Table 1 and needed to computed the geometric gradient and Hessian of p , are obtained from (5.4) and (5.5) as

$$\langle \partial \bar{p}(X), \xi \rangle = \text{tr} \left[d(Y^T Y)^{-1/2} d(Y^T \xi_X Q M) \right] \quad (5.6)$$

and

$$\begin{aligned} \langle \partial^2 \bar{p}(X)[\eta], \xi \rangle = & \text{tr} \left[d(Y^T Y)^{-1/2} d(M^T Q \xi_X^T \eta_X Q M) \right] \\ & - \text{tr} \left[d(Y^T Y)^{-3/2} d(Y^T \eta_X Q M) d(Y^T \xi_X Q M) \right]. \end{aligned} \quad (5.7)$$

Finally, the Riemannian gradient and Hessian of p can be calculated from (5.6) and (5.7) recalling the general formulas (4.7) and (4.9) and the $\mathcal{OB}(2, N)$ -specific expressions given in Table 1. We have therefore all the ingredients to apply Algorithm 4.1 described in Section 4 to find the stationary point of p , i.e., the physical configurations of the laser path. En passant, we mention here that the vector area and compactness ratio of the cavity can be then computed simply as $\mathbf{a}(X; C, Q) = \frac{1}{2} \sum_{k=1}^{N-1} \mathbf{z}_k \wedge \mathbf{z}_{k+1}$, and $\mathbf{k}_r(X; C, Q) = \mathbf{a}(X; C, Q)/p(X; C, Q)$, respectively.

6 Numerical Study

We consider in this section a square optical cavity of side L and four mirrors with same radius r . Correspondingly, the cavity parameters are given by

$$C_{nom} = \left(r - \frac{L}{\sqrt{2}}\right) \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad (6.1)$$

and $Q_{nom} = \text{diagonal}(r, r, r, r)$. The motivation for analyzing such configuration is the design of control algorithms for the GINGER array of ring lasers [8]. In this situation, the optical cavity is made by spherical mirrors whose centers approximately lie on a planar square.

The physical configuration for this symmetry ring laser cavity can be computed by hand (and can be used as one of the tests to validate the correctness of the implementation of the algorithm presented in the previous section). The physical configuration is

$$X^* = -(r - L/\sqrt{2})^{-1} C_{nom} \in \mathcal{OB}(2, 4) \quad (6.2)$$

and the corresponding laser path length is, by construction, $p(X^*; C_{nom}, Q_{nom}) = 4L$. It is interesting to compute the eigenvalues of $\text{Hess } p$ at the point X^* . Using the tools that we have created, one verifies that the spectrum of the geometric Hessian for this special configuration and parameters is

$$-\sqrt{2}r \left(1, 1, 1 - \frac{L}{\sqrt{2}r}, 1 - \frac{L}{\sqrt{2}r}, 1 - \frac{\sqrt{2}L}{r}, 1 - \frac{\sqrt{2}L}{r}, 1 - \frac{\sqrt{2}L}{r}, 1 - \frac{4L}{\sqrt{2}r} \right) \quad (6.3)$$

All the eigenvalues of $\text{Hess } p$ are non-zero provided that the ration L/r is not equal to $\sqrt{2}$, $1/\sqrt{2}$, or $\sqrt{2}/4$ which correspond to unstable optical cavity configurations [7]. For a stable optical cavity the Riemannian Hessian has non-zero eigenvalues, therefore X^* is an isolated root of $\text{grad } p$ and it is a *saddle point* for the length function p , as the spectrum of the Hessian (6.3) has both strictly positive and negative eigenvalues.

Observe that, if a stable optical cavity is slightly misaligned from the nominal center configuration C_{nom} , by continuity arguments, one can conclude that the Riemannian gradient will keep having an isolated root as the eigenvalues of the Riemannian Hessian will be different from zero and that the Newton method is expected to converge.

The proposed geometric algorithm based on Newton's method has been tested by Monte Carlo techniques. Optical cavity configurations are generated starting with mirror positions close to square configuration C_{nom} , with $L = 1.6$ m and represented by $\mathbb{R}^{3 \times 4}$ random matrices whose entries are uniformly distributed over the set $\{|C_{ij} - C_{ij}^*| < \sigma\}$, with σ ranging from $10^{-6}L$ to $10^{-2}L$ with a logarithmic spaced step of $L/10$. The radius matrix was set to $Q = rI_{4 \times 4}$, with $r = 4$ m. The chosen values correspond to the design of the GP2 ring laser [7]. The geometric algorithm has been applied to find the saddle point of the function p , starting from the ideal square configuration given by (6.2). This procedure has been repeated 10^4 times to assess whether mirror displacements are small enough for satisfying the convergence properties of the algorithm. All the execution runs showed no ill-conditioning problems in

the Newton vector computation. In addition, the algorithm took *at most 3* iterations to generate a solution such that $\|\text{grad } p(x)\| < 10^{-12}$ m, achieving quadratic convergence (expected as we are employing Newton's method and the Hessian is not singular at the stationary point). In all the Monte Carlo runs, the computed laser spots positions are saddle points of p .

To better illustrate how the algorithm works, we show in Figure 6.1 the typical behavior of $\|\text{grad } p(x)\|$ in a run with $\sigma \sim 0.5$ m, a value much larger than those considered in the previous simulations. The algorithm took 5 iterations to converge. In Figure 6.2 two comparisons are displayed, between the function p and its second order geometric model, and between the function h and the Armijo condition for h , respectively.

We note that, at each Newton iteration, the geometric model better approximates p ; in fact, from the third iteration the corresponding points in Figure 6.2 overlap. Therefore, at each Newton iteration, the Armijo threshold condition is fulfilled in a smaller number of line search iterations. As a consequence, the Armijo line search stops at its first iteration from the third Newton iteration. In the same figure, $m_f(\lambda) = f(x_k) + \langle \text{grad } f(x_k), \lambda \eta_x \rangle + 1/2 \langle \text{Hess } f(x_k)[\lambda \eta_x], \lambda \eta_x \rangle$, $m_h(\lambda) = h(x) + \sigma \langle \text{grad } h(x_k), \lambda \eta_x \rangle$, $\lambda \in [0, 1]$, $\sigma = 1/2$ and $\eta_x = \eta_{x_k}$ is the Newton vector at the iteration k . In Figure 6.2, each of the 5 iterations of the run reported in Figure 6.1 is displayed. Note how the optical path length p is well modeled by the second order approximation $m_f(\lambda)$ almost from the first iteration and this allows for step sizes t_k to be close to one from the beginning as in a pure Newton method without the line

search condition (4.3). Clearly, this is due also by the fact that we provide the algorithm with a good initial condition.

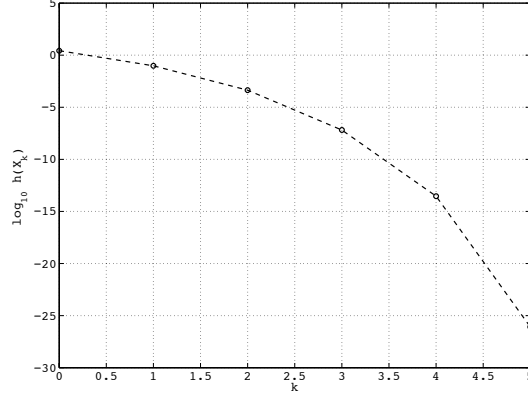


Fig. 6.1 Plot of $h(x_k) = \|\nabla p(x_k)\|$ versus the iteration index k . The data refers to a simulation run with $\sigma \sim 0.5$ m. The algorithm converged in just 5 iterations.

7 Conclusions

We have addressed the problem of computing the light path in optical cavities as a function of the mirror positions and orientations by means of the Fermat's principle. To find the stationary optical path in a polygonal cavity, a geometric algorithm based on Newton method is proposed, exploiting the embedding of the Oblique Manifold $\mathcal{OB}(2, N)$ in the Euclidean space $\mathbb{R}^{3 \times N}$ for the computation of Riemannian gradient and Hessian of the optical path length. The Riemannian gradient and Hessian are then used in each iteration of the algorithm. The algorithm exhibits second order convergence rate.

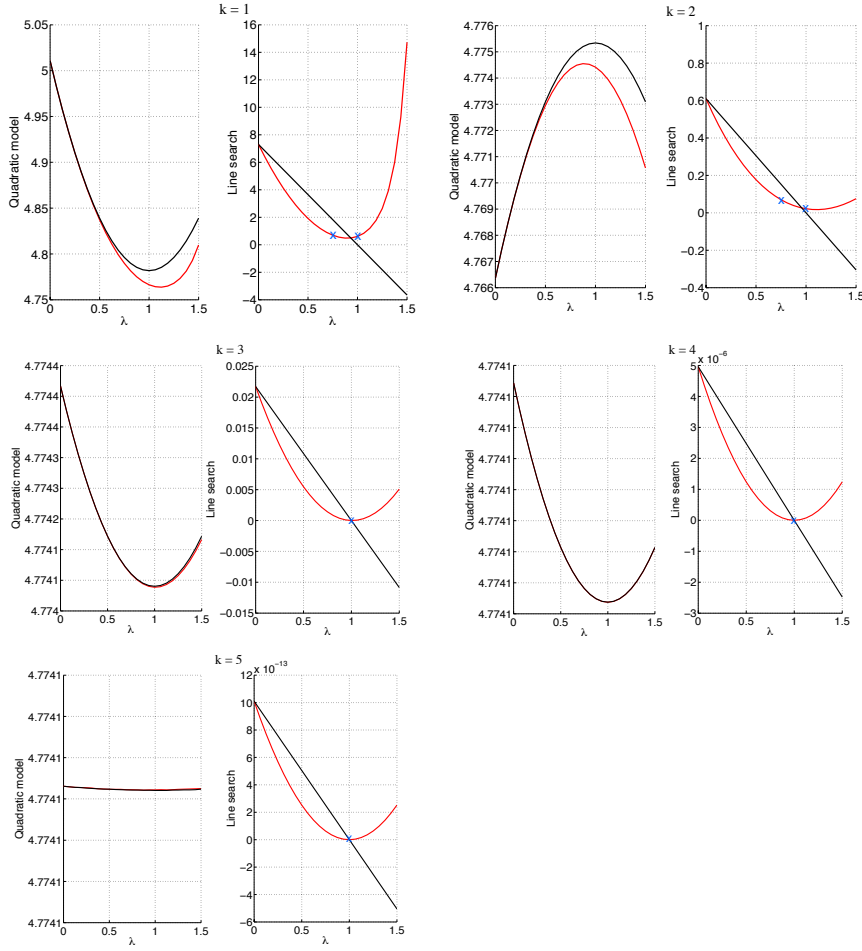


Fig. 6.2 a): Comparison between the functions p (red lines) and its second order geometric model m_f (black lines). b): Comparison between the function h (red lines) and m_h , related to the Armijo condition on h (black lines). The blue crosses represent the iterates of the line search method. The plots are relative to consecutive algorithm iterations from 1 to 5.

Our approach is motivated by the need to compute the optical path of the laser beams in resonant cavities and the algorithm has the potential to allow for the application of control techniques to constrain the beams path.

We showed that the distance function meets the regularity requirements in a neighborhood of the solution for the perfectly aligned square cavity. Monte Carlo simulations have shown that this neighborhood contains all the cavity configurations that can be encountered in practice for heterolithic square ring lasers. In fact, our algorithm has been reported to evaluate the beam steering of a square cavity even with mirror positioning errors $\|C - C^*\| 10^{-3}$ m, while precision machinery of ring laser frames ensure mirrors positioning within $\sim 10^{-5}$ m.

The proposed geometric algorithm provides a relative accuracy of 1 part in 10^{16} in evaluating the optical cavity configuration of a square ring laser. It is worth noticing that greater precisions can be achieved, even if they are not of physical interest. The computational cost of the proposed method is very low, since at most 3 iterations are required to reach the desired accuracy in the Riemannian gradient norm in Monte Carlo simulations.

The geometric approach described in the present paper seems to be well suited to deal with geometrical optics problems where Gaussian profiles are used for beam description. Based on the presented algorithm, future work will be devoted to the calibration and active control of resonant optical cavities [25].

References

1. K.U. Schreiber and J-P. R. Wells, Large ring lasers for rotation sensing, *Rev. Sci. Instrum.* 84, 4, 041101, (2013).

2. D. Cuccato, et al. Controlling the non-linear intracavity dynamics of large He- Ne laser gyroscopes, *Metrologia*, 51, 1, 97, (2014).
3. A. Beghi, et al. Compensation of the laser parameter fluctuations in large ring-laser gyros: a Kalman filter approach, *Appl. Opt.* 51, 7518-7528, (2012).
4. K. U. Schreiber, A. Gebauer, and J.-P. R. Wells, Long-term frequency stabilization of a 16 m^2 ring laser gyroscope, *Opt. Lett.* 37, 1925-1927 (2012).
5. F. Bosi, et al. Measuring gravitomagnetic effects by a multi-ring-laser gyroscope, *Phys. Rev. D*, 84, 12, 122002, (2011).
6. A. Di Virgilio et al. Performances of G-Pisa: a middle size gyrolaser, *Classical and Quantum Gravity*, 27, 8, 084033, (2010).
7. R. Santagata et al. Optimization of the geometrical stability in square ring laser gyroscopes, *Classical and Quantum Gravity*, 32, 5, 055013, (2015).
8. J. Belfi et. al. Interferometric length metrology for the dimensional control of ultra-stable ring laser gyroscopes, *Classical and Quantum Gravity*, 31, 22, 225003 (2014).
9. A. Di Virgilio et al. A ring lasers array for fundamental physics, *C. R. Physique*, 15, 10, 866-874 (2014).
10. For the Romy project see: <http://www.geophysik.uni-muenchen.de/ROMY/>.
11. Y. J. L. Xingwu and C. Meixiong, Generalized ray matrices for spherical mirror reflection, *Optics Express* 19, 7, 6762-76, (2011).
12. B. E. A. Saleh and M. Carl, *Resonator Optics*, John Wiley & Sons, Inc. Teich (2001).
13. P.-A. Absil, R. Mahony and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*, Princeton University Press, (2008).
14. J. E. Marsden and T. Ratiu, *Manifolds, Tensor Analysis, and Applications*, Springer, (2002).
15. Ghatak, Ajoy (2009), *Optics* (4th ed.), ISBN 0-07-338048-2.
16. M. Ishteva, et. al., Differential-geometric Newton method for the best rank- (R_1, R_2, R_3) approximation of tensors, *Numerical Algorithms*, 51, 2, 179-194 (2009).
17. M. Kallay, A geometric Newton-Raphson strategy, *Computer Aided Geometric Design*, 18, 8, 797-803 (2001).

18. L. De Lathauwer, B. De Moor, and J. Vandewalle, On the best rank-1 and rank- (r_1, r_2, \dots, r_n) approximation of higher-order tensors, SIAM. J. Matrix Anal. & Appl., 21(4), 1324-1342 (2000).
19. S. Amat, S. Busquiera and J.M. Gutierrez, Geometric constructions of iterative functions to solve nonlinear equations, Journal of Computational and Applied Mathematics, 157, 1, 197-205 (2003).
20. P.-A. Absil, M. Ishteva and L. De Lathauwer, A Geometric Newton Method for Oja's Vector Field, MIT press journal on Neural Computation, 21, 5, 1415-1433 (2009).
21. P.-A. Absil, R. Mahony, J. Trumpf, An Extrinsic Look at the Riemannian Hessian, Geometric Science of Information, 8085, 361-368, (2013).
22. P.-A. Absil and K.A. Gallivan. "Joint diagonalization on the oblique manifold for independent component analysis." Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on. Vol. 5. IEEE, 2006.
23. L. Armijo, Minimization of functions having Lipschitz continuous first partial derivatives, Pacific Journal of Mathematics 16, 1, 1-3 (1966).
24. Nocedal, Wright, Numerical Optimization, 2006, ISBN:978-0-387-30303-1.
25. A. Beghi, D. Cuccato, A. Ortolan, A. Saccon, Shape and pose of four points in \mathbb{R}^3 (in preparation).

A Discussion and Convergence of Algorithm 1

In Algorithm 1 at each iteration the Newton equation (4.1) is solved for the function f , then the function $h(x) = \|\text{grad } f(x)\|^2$ is minimized along the computed direction. In this way we need to compute only $\text{Hess } f$ and $\text{grad } f$, avoiding the computation of $\text{Hess } h$, that would require to compute the third derivative of f .

Proposition A.1 *Algorithm 1 converges to the stationary point x^* of the function f with quadratic convergence rate, provided that, in a neighborhood $\mathcal{I}(x^*)$ of x^* , $\text{grad } f \neq 0$, $\text{Hess } f$ is injective, and the first iterate is $x_0 \in \mathcal{I}(x^*)$.*

Proof Let x denote a generic algorithm iterate, By hypotheses the Newton vector η_x , solution of (4.1), is well defined. The Riemannian gradient of h reads

$$\text{grad } h = 2\text{Hess } f[\text{grad } f] . \quad (\text{A.1})$$

By evaluating the expression $Dh(x)[\eta_x]$ we get

$$\begin{aligned} Dh(x)[\eta_x] &= 2 \langle \text{grad } f(x), \text{Hess } f(x)[\text{Hess } f(x)^{-1}[-\text{grad } f(x)]] \rangle \\ &= -2 \|\text{grad } f(x)\|^2 \\ &= -2h(x). \end{aligned} \quad (\text{A.2})$$

The sequence $\{\eta_{x_k}\}$ is gradient related to $\{x_k\}$. In fact by hypothesis and (A.1) it holds $\text{grad } h(x_k) \neq 0$, therefore, using (A.2) we get $-2 \sup_{\mathcal{I}(x^*)} h(x_k) = \sup_{\mathcal{I}(x^*)} Dh(x_k)[\eta_{x_k}] < 0$. By the smoothness of the functional $\text{Hess } f$ and of the vector field $\text{grad } f$, since $\mathcal{I}(x^*)$ is a compact set, we can conclude that $\{\eta_{x_k}\}$ is bounded. Hence Algorithm 1 fits in the framework of Theorem 4.3.1 and Theorem 6.3.2 [13, Chs.4-6], stating that every accumulation point of $\{x_k\}$ is a critical point of h , so that the local quadratic convergence holds. \square

Note that the Armijo condition (4.3) for the function h and the direction η_x can be rewritten as

$$h(x) - h(y_k) < -\sigma\gamma_k Dh(x)[\eta_x] = 2\sigma\gamma_k h(x) \quad (\text{A.3})$$

$$h(y_k) > (1 - 2\sigma\gamma_k) h(x) , \quad (\text{A.4})$$

where $y_k = R_x(\gamma_k \eta_x)$, $x = x_k$, $\eta_x = \eta_{x_k}$, and k is the iteration number.